(12) UK Patent Application (19) GB (11) 2 180 539 (13) A

(43) Application published 1 Apr 1987

(21) Application No 8618123

(22) Date of filing 24 Jul 1986

(30) Priority data
(31) 8518753 (32) 24 Jul 1985 (33) GB

(71) Applicant
Glaxo Group Limited

(Incorporated in United Kingdom)

Clarges House, 6–12 Clarges Street, London W1Y 8DH

(72) Inventors
Adrian Hobden
Colin Dykes

(74) Agent and/or Address for Service
Frank B Dehn & Co.
Imperial House, 15–19 Kingsway, London WC2B 6UZ

(51) INT CL⁴
C12N 15/00 C07K 13/00 C12P 21/02

(52) Domestic classification (Edition I):
C3H 672 690 B7M
C6Y 125 501 503
U1S 2413 2415 C3H

(56) Documents cited
None

(58) Field of search
C3H
Selected US specifications from IPC sub-classes C12N
C07G

(54) Microbiological products

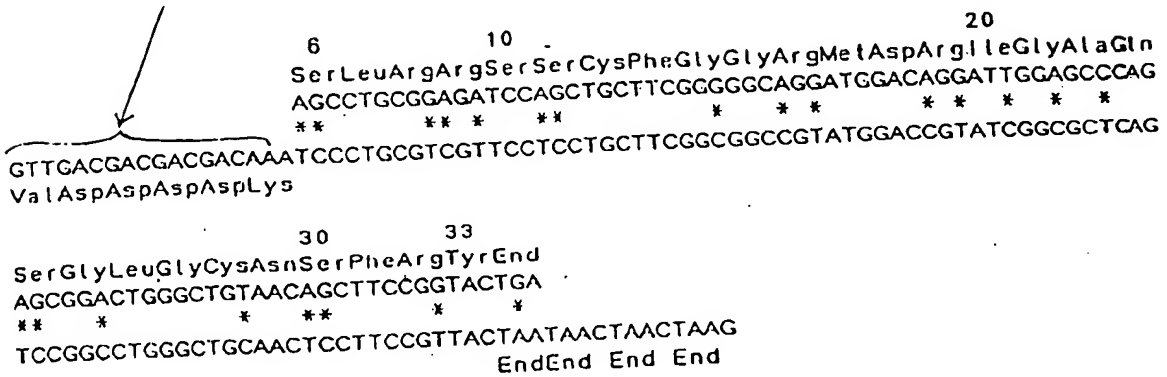
(57) A hybrid protein is provided comprising a first polypeptide having human ANF activity and a carrier polypeptide separated from said first polypeptide by a linker polypeptide containing a recognition site for a proteolytic enzyme. The hybrid protein avoids the degradation of the short chain human ANF polypeptide by the proteases of the transformed host cell. e.g. *E. COLI.*

The enzyme may be enterokinase, thrombin, plasmin collagenase, *Staph, aureus* V8 protease, Factor Xa or endopeptidase lys C.

GB 2 180 539 A

Fig 1

a) Coding sequence for the enterokinase recognition site



b) Coding sequence for the S. aureus V8 protease recognition site

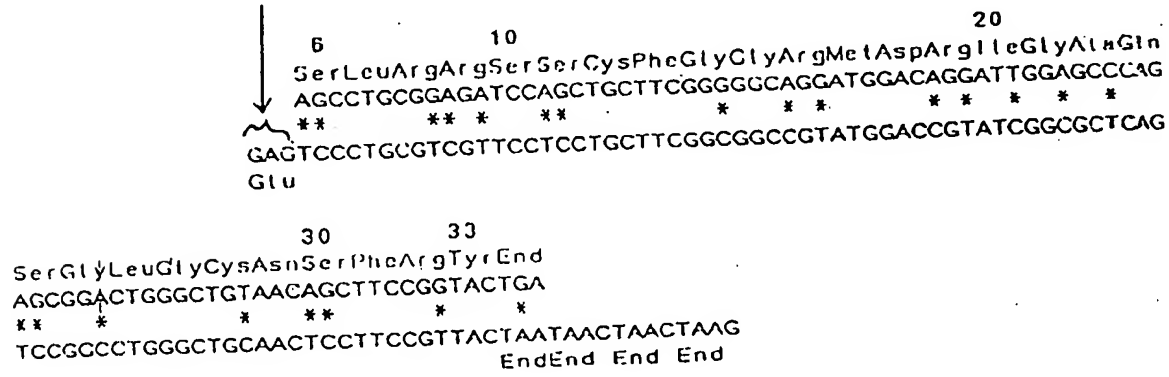
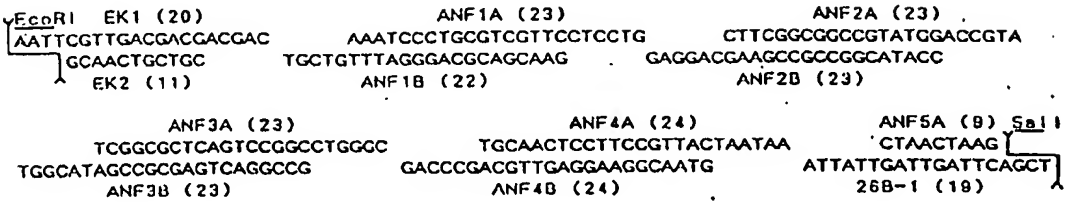


Fig 2 - constructs showing division into individual oligomers.

(1) Constructs for cleavage by enterokinase.

a) ANF-Eco.



b) ANF-Nco.

c) ANF-Sca.

(2) Construct for cleavage by S.Aureus V8 protease.

a) ANF-Xba.

Figure 3

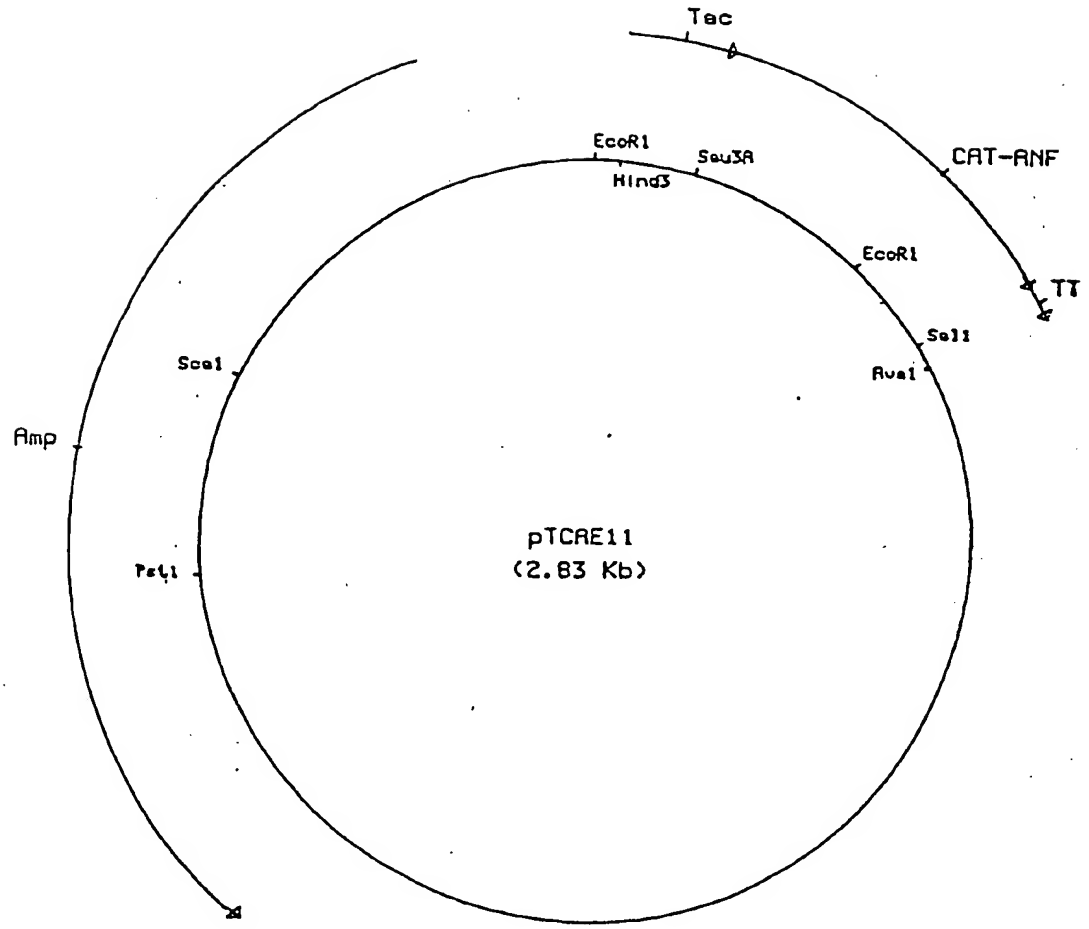

Tac
EcoR1
Sau3A
Hind3
CAT-ANF
EcoR1
TT
Sal1
Ava1
Sca1
Amp
pTCAE11
(2.83 Kb)
Pst1

Figure 4

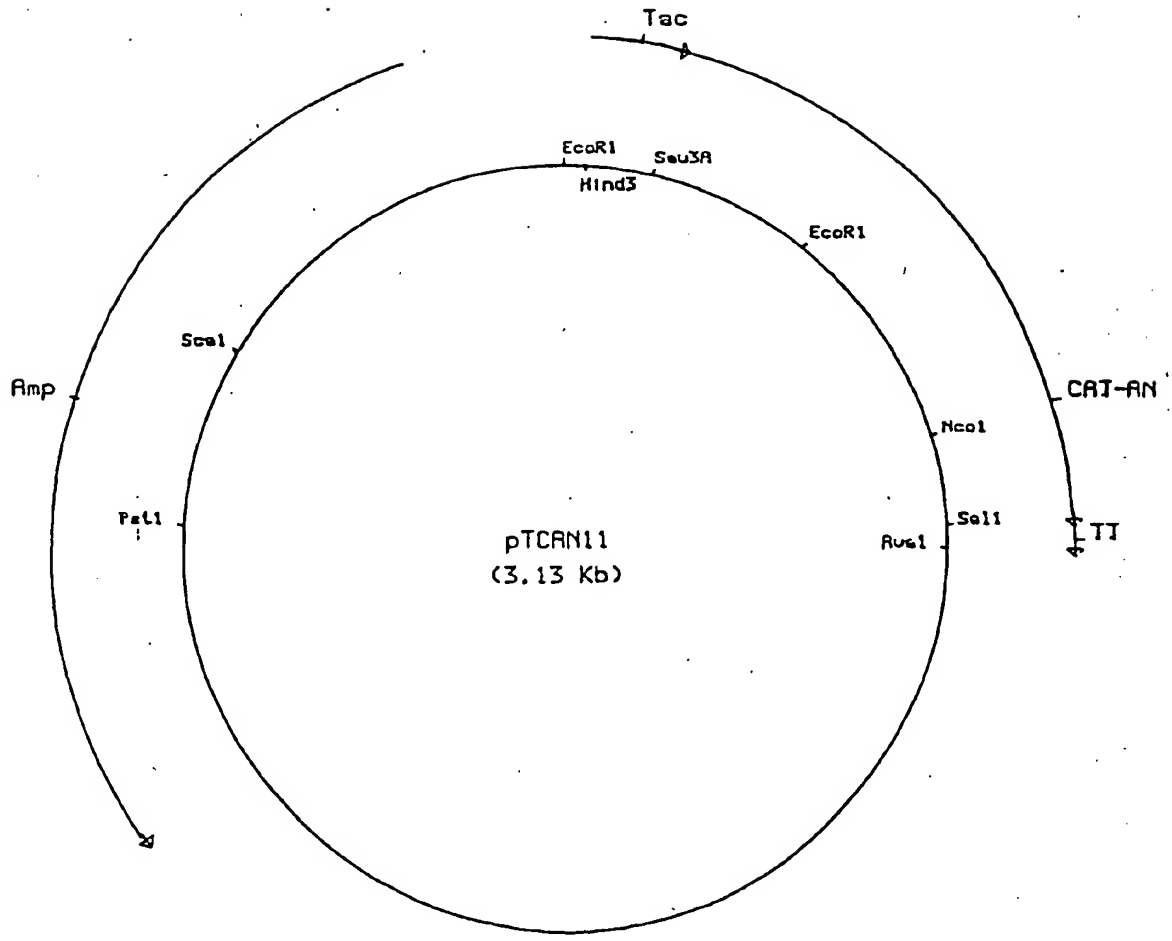

Tac
EcoR1
Sau3A
Hind3
EcoR1
ScaI
Amp
CAT-AN
NcoI
PstI
SalI
RvuI
TI
pTCAN11
(3.13 Kb)

Figure 5

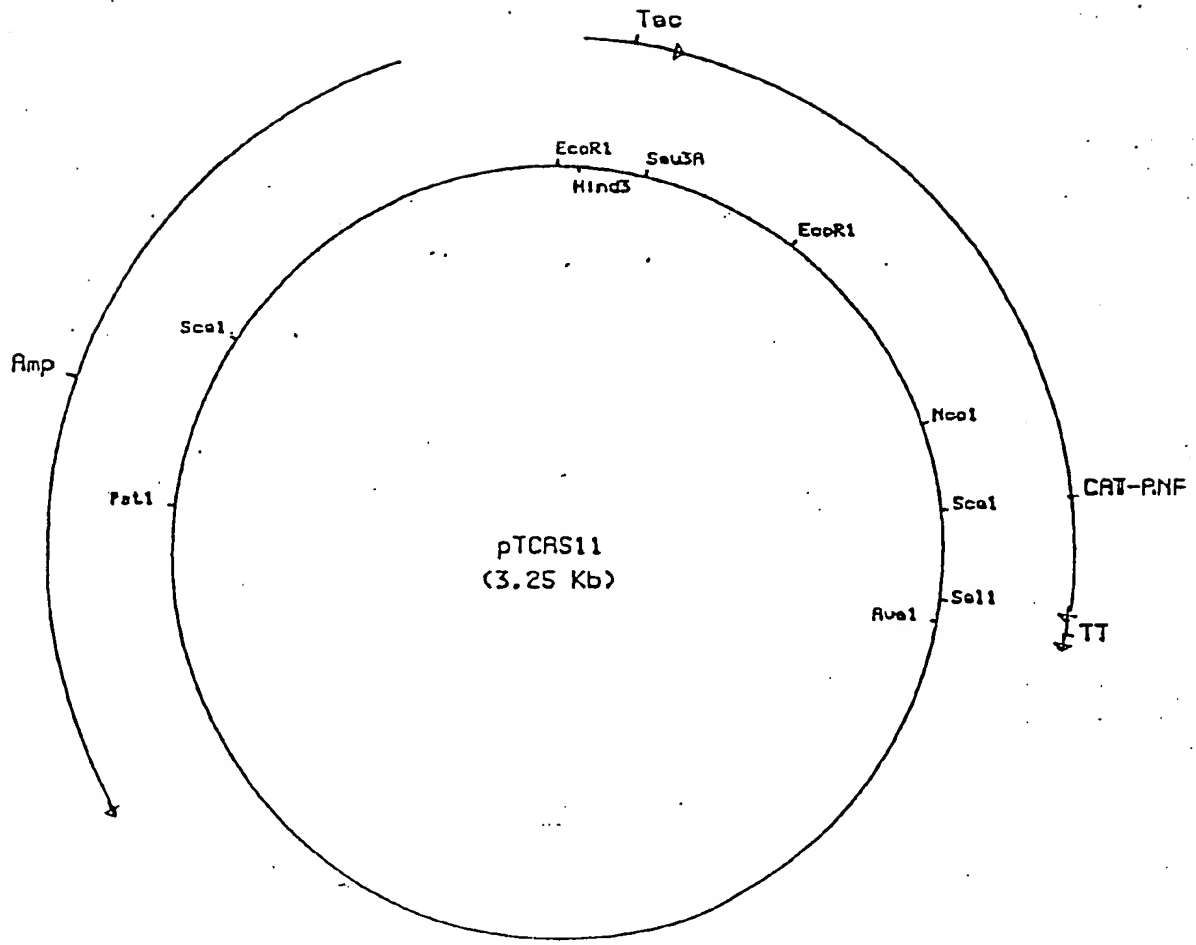

Tac
EcoR1
Sau3A
HindIII
EcoR1
ScaI
Amp
NcoI
PstI
ScaI
CAT-PNF
pTCAS11
(3.25 Kb)
AvaI
SalI
TT

Figure 6

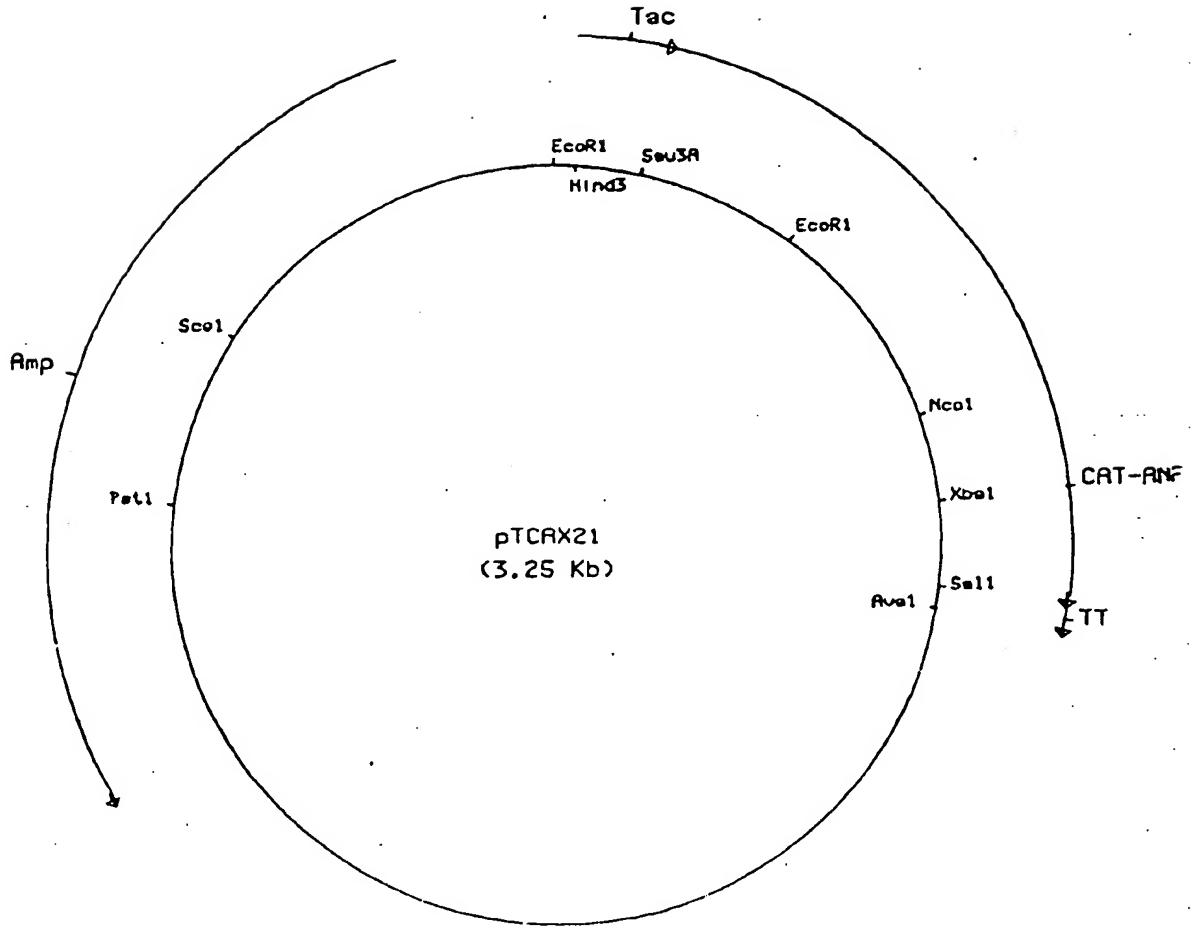
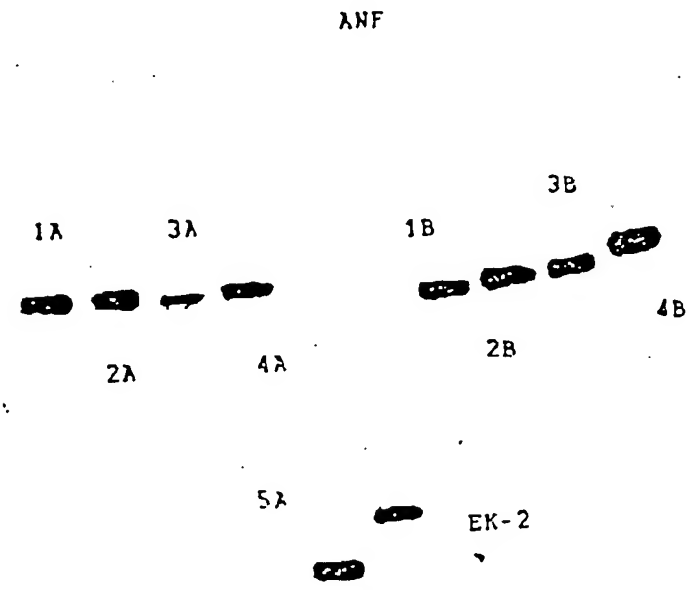

Tac
EcoR1
Sau3A
Hind3
EcoR1
Sca1
Amp
Ncol
Pst1
Xba1
CAT-ANF
pTCAX21
(3.25 Kb)
Ava1
Sal1
TT

FIG 7 PHOTOCOPY OF AN AUTORADIOGRAM OF A 20% POLYACRYLAMIDE SIZING GEL OF THE 10 RADIOLABELLED OLIGOMERS REQUIRED FOR THE LIGATION OF THE ANF-Eco CONSTRUCT.



SPECIFICATION

Microbiological products

The present invention relates to microbiological products and more particularly is concerned with synthetic composite structural genes coding for polypeptides having human atrial natriuretic factor (hereinafter referred to as human ANF) activity and for hybrid proteins containing such polypeptides having human ANF activity as well as with expression vector systems for the production of such hybrid proteins and with the conversion thereof to polypeptides having human ANF activity.

Human atrial natriuretic factor is a polypeptide of 28 amino acid residues which can be found in the artria of the human heart and is believed to be involved in the regulation of fluid volume and electrolyte homeostasis. A similar polypeptide was first identified in extracts of rat heart atria and this polypeptide possesses potent vasodilator activity and inhibits production of aldosterone, the principal mineralcorticoid involved in fluid and electrolyte balance. Atrial natriuretic factor has since been found to be present in cardiac atria of a number of mammalian species.

Recently, there has been considerable interest in the manner of biosynthesis of ANF *in vivo*. It is known that in mammalian cardiac atria ANF is derived from a larger precursor protein containing the polypeptide cardiodilatin which possesses potent vasorelaxant activity. ANF is at the C-terminus of this precursor protein and cardiodilatin immediately follows the signal peptide at the N-terminus. The sequences of cDNAs encoding the precursor proteins containing rat and human ANF have been elucidated [Nakayama, K. et al. (1984) Nature *310*, 699–701].

To permit further study and possible therapeutic use of human ANF and related polypeptides having the same or similar activity, there is a need for a convenient means for preparing such polypeptides in high yield. The present invention provides expression vector systems coding for polypeptides having human ANF activity to enable their production in high yield, for example by a suitably transformed microorganism.

The term microorganism as used herein is intended to include any cells which may be cultured using a suitable nutrient medium and includes, in particular, bacteria such as E. coli and fungi such as yeasts.

Microorganisms, e.g. *E. coli* transformed by a suitable expression vector system carrying a structural gene coding for human ANF alone may not exhibit high production of human ANF, even if the structural gene and the vector system are optimised for expression of the polypeptide. This is due to the fact that relatively short polypeptides such as human ANF can be rapidly degraded by host cell proteases. We have overcome this problem by constructing double-stranded polydeoxyribonucleotides which code for a hybrid protein consisting of a polypeptide having human ANF activity and a carrier polypeptide separated by a linker polypeptide having a recognition site for a proteolytic enzyme which enables the polypeptide having human ANF activity subsequently to be released. Preferably, the enzyme should be a sequence-specific proteolytic enzyme (i.e. one which recognises the above recognition site, but does not cleave at any site in the desired polypeptide) which is absent in the intended host cells. In such hybrid proteins, the human ANF is much more resistant to host cell proteases.

According to one aspect of the present invention, we therefore provide a hybrid protein comprising a first polypeptide having human ANF activity and a carrier polypeptide separated from said first polypeptide by a linker polypeptide containing a recognition site for a proteolytic enzyme.

When the hybrid protein according to the invention is cleaved by the said proteolytic enzyme, a polypeptide having human ANF activity will be released. Such polypeptides, insofar as they differ in their amino acid sequence from natural human ANF are new and constitute a feature of the invention. In particular, if the linker polypeptide is such that cleavage by the proteolytic enzyme does not take place directly at the N-terminal of the human ANF amino acid sequence, there will be a small number of amino acids attached at that end of the molecule.

According to a further feature of the invention we provide a double-stranded polydeoxyribonucleotide coding for a hybrid protein according to the invention wherein the coding strand comprises (i) a sequence coding for a polypeptide having human ANF activity; (ii) fused with the 5'end of sequence (i), a sequence coding for a linker polypeptide containing a recognition site for a proteolytic enzyme; (iii) fused with the 5'end of sequence (ii), a sequence coding for a carrier polypeptide which sequence begins with a translation initiation codon; and (iv) at least one translation stop codon fused with the 3'end of sequence (i).

For convenience, unless otherwise indicated, the expression "human ANF" as used hereinafter includes all polypeptides having human ANF activity.

Conventionally, a double-stranded polydeoxyribonucleotide according to the present invention will possess restriction endonuclease termini such that it can be directly inserted into an appropriate vector system. Preferably, 4 translation stop codons immediately follow the 3' end of the human ANF coding sequence, two fused directly with the human ANF coding sequence in the

normal reading frame and one translation stop codon in each of the two alternative reading frames.

For incorporation into certain expression vector systems, it may be advantageous to construct a double-stranded polydeoxyribonucleotide as hereinbefore described wherein the coding strand further comprises a transcription terminator following the 3' end of the human ANF coding sequence. The chosen transcription terminator preferably comprises at least a 7-mer reverse complementary repeated sequence capable of forming a hairpin loop. It is particularly preferred to use the *trp a* attenuator (which comprises a 7-mer reverse complementary repeated sequence) as the transcription terminator.

The carrier polypeptide coding sequence beginning with the translation initiation codon and fused with the 5' end of the linker sequence is desirably much longer than the human ANF coding sequence, preferably in the range 50 to 300 codons. As indicated above, there will be an adjacent sequence coding for a linker polypeptide containing a recognition site for a proteolytic enzyme, for example, the mammalian gut proteolytic enzyme enterokinase, thrombin, plasmin, collagenase, *Staphylococcal aureus* V8 protease, Factor Xa or endopeptidase lys C. Where it is desired to produce synthetic human ANF identical with natural human ANF the coding sequence for the recognition site of the chosen proteolytic enzyme should be directly fused with the 5' end of the human ANF coding sequence.

In selecting the carrier polypeptide coding sequence, it is particularly convenient to select a well-characterised natural structural gene which codes for a known protein and to use sufficient of the amino acid coding sequence, starting from the translation initiation codon, that the human ANF-containing hybrid protein coded for by the complete composite structural gene can be isolated and identified by some, at least, of the usual techniques employed to isolate and identify the said known protein.

The protein-coding natural gene sequences for the carrier polypeptide which are most favoured have preferably previously been cloned and have one or more restriction enzyme sites which enable a substantial proportion of the amino acid coding sequence to be readily isolated from appropriate plasmid DNA. They are preferably gene sequences which are well expressed by the intended host micro-organism, e.g. *E. coli.*

Double-stranded polydeoxyribonucleotides according to the present invention which we have found to be particularly suitable for incorporation into *E. coli* expression vector systems may, for example, conveniently comprise largely the first 219, 519 or 636 base pairs of the chloramphenicol acetyl transferase (CAT) structural gene present, for example in the known plasmid pBR 325. Each of these 3 nucleotide sequences ends with a restriction enzyme site. The first 219 base pairs of the CAT gene end in an *Eco*R1 site, the first 519 base pairs end in an *Nco*1 site and the first 636 base pairs end in a *Sca*1 site. If desired, a small oligonucleotide adaptor can be inserted at any of these restriction sites, e.g. *Sca*1, in order to introduce different restriction sites, e.g. *Xba*1 and *Xho*1. Addition of such an adaptor necessarily involves the addition of one or more extra amino acids to the CAT sequence.

As a result of the degeneracy of the genetic code, it is possible to predict numerous alternative sequences which can code for human ANF. Certain of these sequences can be expressed more effectively in a microorganism than the natural coding sequence under identical conditions. We have designed a synthetic structural human ANF gene for incorporation into non-mammalian expression vector systems, especially *E. coli* expression vector systems. Fig. 1 shows the sequence of this structural gene (lower strand) in comparison with the natural human ANF coding sequence (upper strand). It can be seen that our synthetic sequence differs at the 15 codon positions marked with an asterisk. It should be noted that Fig. 1 also shows, attached to the 5' end of the synthetic human ANF coding sequence, two different linker sequences comprising the coding sequence for an enterokinase recognition site [a)] and a *S. aureus* V8 protease recognition site [b)].

In selecting the codons for our preferred structural human ANF gene, we firstly made use of known information regarding the frequency of codon usage in various genes highly expressed in *E. coli* and the availability of t-RNA species within *E. coli.* From this information, it is possible to deduce 'favoured codons' of *E. coli* for each amino acid and determine a sequence of such codons coding for human ANF. The sequence of 'favoured codons' was further modified on the basis of the following criteria:

(a) the need to avoid any restriction sites within the sequence which would interfere with the proposed procedures for incorporating the structural gene into expression vector systems,

(b) the requirements of the chosen method of synthesis by mixing and ligating short oligonucleotide chains with single-stranded ends and

(c) the elimination of potentially troublesome complementary sequences within each strand capable of giving rise to messenger RNA secondary structure.

We believe that the structural human ANF gene illustrated in Fig. 1, which has no reverse complementary sequences within the individual strands able to give rise to transcription or translation difficulties, has substantially an optimum sequence for expression in *E. coli* and for

construction as outlined in (b) above.

The linker sequence preferably also has a high proportion of 'favoured codons' of the intended host consistent with this sequence together with the structural human ANF gene having no reverse complementary sequences within the individual strands capable of forming secondary structure and no restriction sites being present which would interfere with incorporation of the entire double-stranded polydeoxyribonucleotide into chosen vector systems. It is also desirable that the linker sequence may be readily constructed together with the structural human ANF gene by the same strategy as hereinbefore outlined for the construction of the preferred structural human ANF gene alone.

Particularly preferred double-stranded polydeoxyribonucleotides according to the present invention comprise a structural human ANF gene sequence at least 75% of the codons of which are 'favoured codons' for *E. coli*, part of a structural gene coding for the carrier polypeptide CAT, which gene is well expressed by *E. coli* in nature, and a short linker sequence separating the human ANF structural gene and the natural carrier polypeptide gene sequence, the coding strand of which comprises the coding sequence of a proteolytic enzyme recognition site, preferably an enterokinase or *S.aureus* V8 Protease recognition site.

Preferred double-stranded polydeoxyribonucleotides according to the present invention which are especially suitable for *E. coli* expression vector systems contain the structural human ANF gene fused with either the enterokinase recognition site or the *S.aureus* V8 protease recognition site coding sequence as shown in Fig. 1. These codes for hybrid proteins wherein the recognition amino acid sequence for enterokinase (Val-$(Asp)_4$-Lys-) or *S.aureus* V8 Protease (Glu) is fused directly with the N-terminus of human ANF.

We have constructed 3 double-stranded polydeoxyribonucleotides according to the invention wherein the coding strand comprises the human ANF/enterokinase recognition site coding sequence shown in Fig. 1(a), four translation stop codons fused with the 3' end of the human ANF coding sequence and, fused with the 5'- end of the enterokinase recognition site coding sequence, the first 73, 173 or 212 codons of the CAT structural gene. We have further constructed a double-stranded polydeoxyribonucleotide according to the invention wherein the coding strand comprises the human ANF/*S.aureus* V8 protease recognition site coding sequence shown in Fig. 1(b), four translation stop codons fused with the 3' end of the human ANF coding sequence and the first 212 codons of the CAT structural gene plus an additional codon resulting from use in the construction process of an *Xba*1 restriction site on an oligonucleotide adaptor. These 4 double-stranded polydeoxyribonucleotides were constructed so that an additional conveniently prepared fragment comprising a *trp a* attenuator sequence may be readily used as transcription terminator. The resulting longer double-stranded polydeoxyribonucleotides are particularly preferred, for example, for incorporation into expression vector systems based on plasmid pAT153.

Any desired double-stranded polydeoxyribonucleotide according to the present invention may be constructed, by assembling appropriate deoxyribonucleotides or oligodeoxyribonucleotide sections. In a preferred method for preparing a polydeoxyribonucleotide coding for a hybrid protein according to the invention double stranded polydeoxyribonucleotides fragments, with appropriate restriction enzyme termini, are mixed and ligated; the fragments are as follows: (1) a fragment wherein the coding strand comprises the coding sequence for a polypeptide having human ANF activity and, fused with the 3'-end and 5'-end thereof respectively, at least one translation stop codon and a sequence coding for a linker polypeptide containing a recognition site for a proteolytic enzyme, (2) a fragment comprising a sequence coding for the carrier polypeptide and, if appropriate, (3) a fragment which comprises a transcription terminator sequence.

It is difficult to synthesise long single stranded DNA sequencies and it is therefore advantageous to build up the gene coding for human ANF by annealing and ligating relatively short (e.g. 9–24 bases) oligonucloetide sections of the coding strand of the intended gene fragment together with overlapping oligonucleotide sections of the complementary strand. In order to permit accurate annealing, the oligonucleotides within each fragment should not contain complementary sequences which could lead to incorrect annealing.

It is thus preferable for the structural gene coding for human ANF to be designed so that it can be built up from oligonucleotide chains, for example of about 20 bases in length, which are capable of joining in such a way that the oligonucleotides of the coding strand overlap the junctions of the oligonucleotides of the complementary strand by a sufficient number of base pairs to ensure satisfactory annealing to produce the desired structural gene. Each of the overlapping sequences within the gene is unique, so that, after 5'-phosphorylation of the oligonucleotides, with the exception of those at the 5'-ends of each strand of the intended gene, all the oligonucleotides are simply mixed together, followed by annealing and ligation, to produce the desired gene.

Following mixing of appropriate oligonucleotide chains, ligation of the oligonucleotides of each strand is achieved in conventional manner using a DNA ligase such as T_4 DNA ligase, preferably with incorporation of tracer radioactivity so that the length of constructed fragments can be

checked by gel electrophoresis and autoradiography. After the gene has been assembled, it may be cloned in a suitable expression vector and if desired excised from the vector and its structure checked by sequencing.

A shorter fragment comprising a transcription terminator sequence may be readily constructed by the same strategy. Fragment (2) may also be totally synthetic and constructed from smaller oligonucleotide fragments in the same manner. However, if the gene sequence coding for the carrier polypeptide comprises the whole of, or part of, the coding sequence of a natural gene, fragment (2) is preferably a restriction enzyme fragment comprising the desired natural gene sequence derived from appropriate isolated DNA, or a fragment derived from such a restriction enzyme fragment by modification of one or more of the restriction enzyme termini.

Synthesis of individual oligonucleotides for the coding strand (A strand) and complementary strand (B strand) of a fragment which is to be prepared by mixing and ligating appropriate double-stranded oligonucleotide chains as hereinbefore described, can be achieved by such techniques as the phosphotriester and phosphoramidite methods. These procedures are most conveniently carried out on a solid phase support. (M. Gait *et al.*, Chemical and Enzymatic Synthesis of Gene Fragments, ed. H.G. Gassen and A. Long, Verlag Chemie Weinheim 1982; M.D. Matteucci and M.H. Caruthers, J.A.C.S., *103*, 3185–3191 (1981); M.H. Caruthers et al., Genetic Engineering, ed. J. Sethow and A. Hollander, *4*, 1, Plenum Press, New York).

Cloning of each fragment may be carried out by incorporating the fragments into a suitable vector and transforming a suitable host organism such as *E. coli* JM103.

The preferred double-stranded polydeoxyribonucleotides according to the invention code for a hybrid protein consisting of (i) part of the CAT protein; (ii) a polypeptide including the amino acid sequence recognized by enterokinase or *S.aureus* V8 protease and (iii) human ANF. Three fragments for construction of such a preferred polydeoxyribonucleotide comprising the human ANF/enterokinase recognition site coding sequence are shown in Fig. 2. One of these fragments (hereinafter referred to as ANF-Eco), has an *Eco*R1 'cohesive end' plus cytosine-guanosine base pair directly preceding the enterokinase recognition site coding sequence and an *Sal*1 'cohesive end' plus a guanosine-cytosine base pair directly following the final translation stop codon, and is capable of joining with the first part of the CAT structural gene when cut by the restriction enzyme *Eco*R1. A further fragment, which has been designate ANF-Nco and differs from ANF-Eco only in having an *Nco*1 terminus directly preceding the enterokinase site coding sequence, was designed for joining with the first part of the CAT structural gene when cut by the restriction enzyme *Nco*1. The third fragment, which has been designated ANF-Sca, has a *Sca1* terminus preceding the enterokinase recognition sequence. Also shown in Fig. 2 is a fourth fragment, which has been designated ANF-Xba, suitable for construction of a preferred double-stranded polydeoxyribonucleotide according to the invention encoding a hybrid protein with an *S. aureus* V8 protease recognition site. This fragment has a *Xba*1 cohesive end which includes the first base of the codon for the glutamic acid residue serving as the *S.aureus* V8 protease recognition site. The precise oligonucleotide chains from which each of these fragments were built up are indicated in Fig. 2.

Vectors which contain a double-stranded polydeoxyribonucleotide according to the invention, e.g. expression vectors which contain a double-stranded polydeoxyribonucleotide according to the invention at an appropriate site for expression of the composite structural gene, constitute yet another aspect of the present invention. In such an expression vector, the chosen composite structural gene coding in part for human ANF will be within the transcriptional unit of a promoter recognized by the RNA polymerase of the intended host, e.g. *E.coli.* Suitable promoters for *E.coli* expression vector systems, for example, include the *lac* and *trp* promoters of the *E.coli* genome. If the intended host is *E.coli*, it may be particularly preferable to construct an expression vector system according to the invention wherein the chosen human ANF-coding composite structural gene is under the control of a *tac* promoter/operator, (e.g. *tac*-1), which is made up of two elements, one from the *trp* promoter and one from the *lac* promoter/operator region. [De Boer, H. et al. (1983) P.N.A.S. *80*, 21]. Preceding the composite structural gene coding for the human ANF-containing hybrid protein, there should also be an appropriately sited Shine-Dalgarno sequence (i.e. a sequence which when present in a RNA transcript assists binding of ribosomes by interaction with the 3' end of the 16s ribosomal RNA).

We have constructed four vector systems according to the invention which are especially suitable for *E. coli*, all utilising elements of the CAT structural gene obtained from the known plasmid pBR325. The four expression vector systems based on plasmid pAT153 have inserted downstream from a *tac* promoter/operator one of the four preferred double-stranded polydeoxyribonucleotides herein before described comprising part of the CAT structural gene, the ANF-Eco, ANF-Nco, ANF-Sca or ANF-Xba sequence and a *trp a* transcription terminator. The *tac* promoter/operator, between a HindIII restriction site and a BamH1 restriction site, directs transcription of the human ANF-coding composite structural gene.

For the construction of the four vector systems, herein termed pTCAN11, pTCAS11 and pTCAX21, a HindIII-BamH1 restriction enzyme fragment comprising a *tac* promoter was derived

from the known plasmid pDR540 and ligated with *Hind*III/*Bam*H1 digested pAT153 DNA. Hybrid plasmids were selected which confer resistance to ampicillin on *E. coli* cells. The selected hybrid plasmids were then digested with *Sal*1 and *Ava*1 restriction enzymes and ligated with a synthetic *trp a* transcription terminator. A plasmid was selected which carried the terminator and this hybrid was digested with the restriction enzyme *Bam*H1 and ligated with DNA fragments of plasmid pBR325 previously digested with the restriction enzyme *Sau*3A. A hybrid plasmid was selected which conferred resistance to ampicillin and (upon induction with isopropyl β–D–thiogalactoside) resistance to chloramphenicol. If desired, the selected plasmid may be digested with *Sca*1 and a synthetic oligonucleotide adaptor carrying a *Xba*1 recognition site introduced. Selected plasmids may then be digested with *Eco*R1 and *Sal*1, *Nco*1 and *Sal*1, *Sca*1 and *Sal*1 or *Xba*1 and *Sal*1 and the appropriate DNA fragments required to complete the desired expression vector systems may be introduced and ligated. Structural maps of the four preferred expression vector systems are included in the accompanying drawings (Figs. 3, 4, 5 and 6).

The present invention extends to cells transformed by a vector, e.g. an expression vector such as a plasmid as hereinbefore described and to processes for the preparation of human ANF involving the use of such transformed cells.

According to a still further aspect of the present invention, we provide a process for the preparation of human ANF which comprises growing transformed cells containing an expression vector as hereinbefore described under conditions whereby the composite structural gene coding in part for human ANF is expressed, contacting the human ANF-containing hybrid protein thus produced with the appropriate proteolytic enzyme to release the human ANF polypeptide and isolating the desired polypeptide thus released.

In the process for growing the cells, inoculum development may proceed through preparation of ampoules containing transformed cells followed by cell growth on minimal agar slopes. The inoculum may then be used to inoculate a Florence flask containing a minimal medium, and this culture used to initiate cell growth in large fementation vessels. The fermentation medium will normally contain a source of nitrogen, such as ammonium sulphate, a source of carbon and energy such as glucose or glycerol, and trace elements. In the case of *E. coli*, submerged aerobic fermentation is preferred, advantageously at about 37°C.

If in the expression vector employed, the human ANF-coding composite structural gene is under the control of an inducible promoter system, an appropriate inducer may be added to the growth medium. For example, in the case of cells transformed by an expression vector according to the invention which has a *tac* promoter/operator for control of transcription of the human ANF-coding composite structural gene, production of the human ANF-containing hybrid protein may be induced by addition to the growth medium of isopropyl-β-D-thiogalactoside (IPTG).

The hybrid protein may be produced in a soluble form or as granular inclusion bodies which can be recovered, after cell lysis, by differential centrifugation. Soluble proteins may be rcovered by conventional means, e.g. by affinity chromatography. Insoluble proteins can be solubilised by conventional methods, e.g. by addition of sodium dodecyl sulphate, urea, a surfactant, such as Triton X-100, or guanidine hydrochloride, and purified by any known method such as precipitation with ammonium sulfate, dialysis to remove salts (under normal or vacuum pressure), gel filtration, chromatography, preparative flatbed iso-electric focusing, gel electrophoresis, high performance liquid chromatography (hereinafter "HPLC"), ion exchange and affinity chromatography on dye bound carrier, on activated Sepharose 4B coupled with monoclonal antibody against the hybrid protein or on lectin bound Sepharose 4B and the like.

The hybrid protein may be further purified prior to contact with the appropriate proteolytic enzyme, for example by means of conventional methods such as chromatography. The desired hybrid protein may be recognised and assayed by such methods as radioimmunoassay based on the known immunological properties of the carrier polypeptide.

Human ANF released from a hybrid protein according to the present invention by the action of the appropriate proteolytic enzyme may be conveniently isolated by conventional techniques such as those described above for the hybrid protein. The human ANF may be recognised and assayed by bioassay techniques. If desired, assay of the hybrid protein may be effected by cleavage to human ANF and bioassay of the latter.

The following non-limiting examples serve to illustrate the present invention more fully.

Materials

Dithiothreitol (DTT), spermidine, agarose type 1, sodium dodecyl sulphate (SDS), 2-mercaptoethanol, ATP, bovine serum albumin chloramphenicol caproate were from Sigma (London) Chemical Co.

T4 DNA ligase, T4 polynucleotide kinase, *E. coli* DNA Polymerase 1 'Klenow Fragment' and the restriction endonucleases, BamH1, EcoR1, Ava1, Sca1, HindIII, Sau3A and Xba1 were from the Boehringer Corporation (London) Ltd.

The restriction endonuclease Nco1 was from New England Biolabs Ltd. *S. Aureus* V8 protease was from Worthington™. Enterokinase was purified from pig's intestines as described by Bro-

drick *et al* (1978) I. Biol. Chem. *253*, 2737–2742.

Acrylamide and N-N'-methylenebisacrylamide and P_4 Biogel were from Bio-Rad Laboratories Ltd.

DE52 cellulose was supplied by Whatman Chemical Separation Ltd.

øX174 RF DNA-HaeIII and m13mp8 RF DNA were from Bethesda Research Laboratories (UK) Ltd.

dATP, dCTP, dTP, dTTP, ddATP, ddCTP, ddGTP, dTTP, plasmid pDR540 and *E. coli* strain JM103 were from P-L Biochemicals, Division of Pharmacia Inc.

(δ^{32}P)ATP, (δ-thio^{35}S) dATP, plasmid pAT153 and plasmid pBR325 were from Amersham International PLC.

Partisil 10-SAX was supplied by R.O.D. Scientific, (Merseyside).

Biosearch Synthesis Automation Machine (SAM) and Fractosil 500 were supplied by Biosearch Inc. (San Rafael, USA).

Fractosil 500 was functionalised with projected nucleosides purchased from Cruachem Ltd using the procedures described in M H Caruthers in "Chemical and Enzymatic Synthesis of Gene Fragments – A Laboratory Manual", Eds H. G. Gassen and A. Lang, Verlag Chemie, 1982, p71.

The nucleotides used in the coupling reaction were purchased from Cruachem Ltd and were used as appropriate in the form of 5'-dimethoxytritylthymidine-3'-(2-chlorophenyl) phosphate triethyl ammonium salt, 5'-dimethoxytrityl-N^6-benzoyl-2'-deoxyadenosine-3'-(2-chlorophenyl)-phosphate triethyl ammonium salt, 5'-dimethoxytrityl-N^2-isobutyryl-2'-deoxyguanosine-3'-(2-chlorophenyl)-phosphate triethyl ammonium salt or 5'-dimethoxytrityl-N^4-benzoyl-2'-deoxycytidine-3'-(2-chlorophenyl)-phosphate triethylammonium salt.

All solutions were made up in high quality HPLC grade water. The pH of solutions was adjusted by the addition of appropriate acid or base as required at 22°C.

Bacterial Strains and Plasmid

The *E. coli* strain used was JM103. The plasmids used were pDR540, pBR325 and pAT153. The derivation of pAT153 has been described (Twigg, A.J. and Sherrat, D. (1980) Nature *283* pp 216–218). The sequencing 'phage was M13mp8.

All temperatures throughout the Examples are given in degree centigrade.

Abbreviations

MSCl —Mesitylene Sulphonyl Chloride
MeCN —Acetonitrile
DCM —Dichloromethane
DMAP—4,-dimethylaminopyridine
DMTr —4,4-dimethoxytrityl
NMI —N-methylimidazole
THF —Tetrahydrofuran
DTT —Dithiothreitol
BSA —Bovine serum albumin

EXAMPLE 1

TABLE 1 – Synthesis, deprotection and purification methods for oligonucleotides comprising the ANF FUSION constructs

Oligo No.	Sequence	Synthesis deprotection & purification methods
EK1	AATTCGTTGACGACGACGAC	A, C, E
EK1-1	CATGGTTGACGACGACGAC	A, C, D
EK 1-3	ACTGCGTTGACGACGACGAC	A, C*, D
ANF 1A	AAATCCCTGCGTCGTTCCTCCTG	A, C, D
ANF 2A	CTTCGGCGGCCGTATGGACCGTA	A, C*, D
ANF 3A	TCGGCGCTCAGTCCGGCCTGGGC	A, C, E
ANF 4A	TGCAACTCCTTCCGTTACTAATAA	A, D, E
ANF 5A	CTAACTAAG	A, C, E,
27A	TCGACAGCCCGCCTAAT	B, C, E
28A	GAGCGGGCTTTTTTT	B, C, E
EK2	CGTCGTCAACG	A, C*, E
EK2-1	CGTCGTCAAC	A, C, E
EK2-2	CGTCGTCAACGCAGT	A, C*, D
ANF 1B	GAACGACGCAGGGATTTGTCGT	A, C, E
ANF 2B	CCATACGGCCGCCGAAGCAGGAG	A, C, D
ANF 3B	GCCGGACTGAGCGCCGATACGGT	A, C*, D
ANF 4B	GTAACGGAAGGAGTTGCAGCCCAG	A, C*, D
26B-1	TCGACTTAGTTAGTTATTA	A$, C*, E
27B	CGCTCATTAGGCGGGCTG	B, C, E
28B	CCGAAAAAAAAGCC	B, C, E
XX-A	ATCTAGATGCTCGAG	A, C, D
XX-B	CTCGAGCATCTAGAT	A, C, D
ANF 1A-1	CTAGAGTCCCTGCGTCGTTCCTCCTG	A, C, D
ANF 1B-1	GAACGACGCAGGGACT	A, C, D

*As in method C but with pyridine-2-aldoxime instead of syn-2-nitrobenzaldoxime
$ Capping routine omitted.

METHOD A – PHOSPHOTRIESTER PROCEDURE
Synthesis cycle on a SAM machine with silica gel as solid support and MSCl as coupling agent – modified procedure

Step	*Operation*	*Solvents/Regents*	*Time (min)*
1	Loading	Fractosil 500(50mg,2uM)	
2	Wash	MeCN	1.5(flow2.5ml/min)
3	Detritylation	2%-TCA in DCM	1.5(flow2.5ml/min)
4	Wash	MeCN	1.5(flow2.5ml/min)
5	Coupling	Nucleotide(27.5uM)/NMI(1mM) in MeCN-pyridine (7.5:1) MSCl(114uM) in MeCN-pyridine(9:1)	6.5(flow0.7ml/min)
6	Capping	6.6%-Acetic anhydride in THF 3.3%-DMAP in pyridine-THF(1:15)	3 (flow2.5ml/min)
7	Wash	MeCN	5 (flow2.5ml/min)
8	Repeat steps 2–7		
9	Removal	After final coupling, cycle taken to step 7 (ie. DMTr not removed) & support removed & taken through deprotection sequence.	

METHOD B – PHOSPHOTRIESTER PROCEDURE
Synthesis cycle on a SAM machine with silica gel as solid support and MSCl as coupling agent – original procedure

Step	Operation	Solvents/Reagents	Time (min)
1	Loading	Fractosil 500(50mg,2uM)	
2	Wash	MeCN	6 (flow 2.5ml/min)
3	Detritylation	2%-TCA in DCM	2 (flow 2.5ml/min)
4	Wash	MeCN	4 (flow 2.5ml/min)
5	Coupling	Nucleotide (27.5uM)/NMI (1mM) in MeCN-pyridine (7.5:1) MSCl (114uM) in MeCN-pyridine (9:1)	16.5(flow0.7ml/min)
6	Repeat steps 2–5		
7	Removal	After final coupling, cycle taken to step 4 (i.e. DMTr removed) & support removed & taken through deprotection sequence	

METHOD C – Deprotection procedure for oligomers prepared by phosphotriester methodology using syn-2-nitrobenzaldoxime.

1. Dried support in Eppendorf tube treated with 1.0–1.5ml of a solution of syn-2-nitrobenzaldoxime (140mg) & 1,1,3,3-tetramethylguanidine (100μl) in 50%-aq. dioxan (2ml). Mixture shaken and stored for 15–18h at room temperature.
2. Mixture filtered and filtrate evaporated.
3. Residue dissolved in conc. ammonia (5ml) & heated at 60° for 3h in a stoppered flask.
4. Solution cooled then evaporated & residue dissolved in 80%-HOAc (5ml) & left for 30min (NB. acid treatment is omitted if the DMTr group has already been removed in the synthetic cycle, or if reverse-phase HPLC purification is to be employed).
5. Solution evaporated & residue dissolved in water (3ml) & washed with ether (5 × 10ml). Aqueous portion evaporated.
6. Residue purified by HPLC or gel electrophoresis.

METHOD D – Purification of oligomers by preparative gel electro phoresis

1. Crude deprotected oligomer dissolved in 0.3M-NaOAc (300ul) in a 1.5 Eppendorf tube.
2. tRNA (20ug) & EtOH (ca. 1ml) added, mixture vortexed & stored at −78° for 15min and then centrifuged for 10min. Supernatant discarded.
3. Dried pellet dissolved in water (80μl) & portion (40μl) diluted with gel loading solution (EDTA(1mM), tris borate (pH8.3) (50mM), 0.05% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanole & 80% aq. formamide) (40ul).
4. Sample electrophoresed on 20% poyacrylamide gel containing tris borate (pH8.3) (89mM), EDTA (2.5mM) & 7M-urea.
5. Gel visualised by UV & appropriate band excised & incubated with gel elution buffer (0.5M-ammonium acetate, 0.01M-magnesium acetate, 0.1% (w/v) sodium dodecylsulphate & 0.1M-EDTA) (2ml) overnight at 37°.
6. Mixture filtered & process repeated with further buffer (2ml) for 1h at 37°.
7. Combined filtrates washed with n-BuOH until volume ca. 500ul.
8. Solution washed through a P4 Biogel column eluted with EtOH-water (1:4).
9. Eluate containing DNA evaporated.
10. Residue dissolved in known volume of water, solution examined by UV & its concentration adjusted to give a solution of ca. 1 OD/ml.

METHOD E – Purification of oligomers by anion-exchange HPLC

1. Crude deprotected oligomer dissolved in water (500ul).
2. Sample(10μl) examined by anion-exchange HPLC on Partisil-10 SAX column at 40 using a gradient of 0.001–0.3M potassium dihydrogen orthophosphate in formamide-water (3:2) over 45 min & a UV detector at 270nm. (Alternatively a gradient of 0.001–0.4M potassium dihydrogenorthophosphate in water-formamide (3:2) may be employed).
3. Remainder of sample purified in aliquots (up to 200ul) using above gradient or a modified gradient suggested by the analytical run.
4. Eluates corresponding to the product peak combined & desalted on a P4 Biogel column eluted with EtOH-water (1:4).
5. Eluate containing DNA evaporated.
6. Residue dissolved in known volume of water, solution examined by UC and its concentration adjusted to give a solution of ca. 1 OD/ml.

METHOD F – Kination of oligomers for the ANF-Eco construct.

1. The following combined in individual 1.5ml eppitubes

i) A solution containing 100pmoles of the oligonucleotide to be radiolabelled; i.e. ANF1A, ANF2A, ANF3A, ANF4A, ANF5A, EK2, ANF1B, ANF2B, ANF3B and ANF4B.

ii) A solution containing a mixture of both gamma phosphate P32 labelled and cold ATP (200pmoles).

iii) A solution containing kination salts (250mM Tris HCl pH7.6, 50mM $MgCl_2$ 25mM DTT, 5mM spermidine, 5mM EDTA) (4ul).

iv) Distilled, deionised water to a final volume of 20 ul.

v) T4 Polynucleotide kinase (10units BRL).

2. The sample was thoroughly mixed and incubated at 37° for 45mins.

3. An aliquot of each sample (1μl) was diluted with gel loading solution (0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanole in 7M aqueous urea) (5μl).

4. The samples were electrophoresed on a 20% polyacrylamide gel containing tris borate (pH8.3) (89mM), EDTA (2.5mM) and 7M urea.

5. The gel was autoradiographed at −78° and the film developed.

Note:

i) A photocopy of the autoradiogram of the radiolabelled oligomers required to construct ANF-Eco is included as Fig. 6.

ii) The method described above has also been successfully applied to the oligomers required for the ANF-Nco, ANF-Sca and ANF-Xba1 constructs.

METHOD G – Ligation of oligomers comprising the ANF-Eco construct

1. The following combined in a 1.5ml eppitube

i) A solution containing an equimolar amount (44pmoles) of each of the 5' 32P phosphorylated oligomers ANF1A, ANF2A, ANF3A, ANF4A, ANF5A, EK2, ANF1B, ANF2B, ANF3B and ANF4B (220μl).

ii) A solution containing 44pmoles of oligomer EK1 (9.6μl).

iii) A solution containing 44pmoles of oligomer 26B-1 (7.2μl).

iv) A solution containing 20ug tRNA (6.25μl).

v) Distilled, deionised water (56.95μl).

2. Sodium acetate solution (3M 30μl) & EtOH (ca. 1ml) added, mixture vortexed & stored at −78° for 15min and then centrifuged for 10min. Supernatant discarded.

3. Pellet washed with cold (−20°) 80% aqueous ethanol (100μl) and dried in vacuo.

4. Pellet dissolved in ligation salts (50mM Tris pH7.6, 10mM magnesium chloride) (34μl), heated at 60° for 5min and cooled to 23° over 4h.

5. ATP (10mM, 5ul), DTT (200mM, 5μl), BSA (5mg/ml, 5μl), T4 DNA ligase (2units) added and the tube incubated at 23° for 16h.

6. Sample diluted with gel loading solution (50% (w/v) sucrose, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanole) (25μl).

7. Sample electrophoresed on 8% polyacrylamide gel containing tris borate (pH8.3) (89mM) and EDTA (2.5mM).

8. Gel autoradiographed and product band identified by comparison with radiolabelled øX174 RF, DNA *Hae*III Fragment markers.

9. Product excised & incubated with gel elution buffer (0.5M ammonium acetate, 0.01M magnesium acetate, 0.1% (w/v) sodium dodecylsulphate & 0.1M EDTA) (2ml) overnight at 37°.

10. Mixture filtered and process repeated with further buffer (2ml) for 1h at 37°.

11. Combined filtrates diluted with an equal volume of water and sample applied to a DE52 cellulose column.

12. Column washed with 0.05M sodium chloride (2ml).

13. DNA eluted from column by washing with 1M sodium chloride solution, ethanol precipitated in the presence of tRNA (20ug) and the pellet dissolved in distilled, deionised water (20ul).

14. Product quantified by scintillation counting.

Note:

i) This method has also been successfully applied to the assembly of the ANF-Nco, ANF-Sca and ANF-Xba1 constructs from their constitutive oligomers.

ii) Although the gel purified ligation product is the substrate of choice for ligation to a suitably prepared plasmid/phage vector the crude ligation mixture has also been successfully employed.

Example 2

Construction of an Expression Vector for the Production of Fusion Proteins in E. coli

a) *Digestion of pAT153 with HindIII and BamH1*

5μg of pAT153 was dissolved in 50μl of a reaction mixture (final concentration 20mM-Tris/HCl, 10mM-$MgCl_2$, 1mM-DTT, 50mM NaCl, pH 7.5). 10 units of the enzyme *Bam*H1 (10 units/μl) and 10 units of the enzyme HindIII (10 units/μl) were added to the mixture which was

incubated at 37°C for 3 hours resulting in the complete double digestion of the vector DNA. These conditions had been established in a pilot experiment in which the reaction were monitored by agarose-gel electrophoresis as follows:

0.7% agarose slab gel (25×20×0.4cm) was prepared in running buffer (40mM-Tris acetate, 2mM EDTA, pH 7.8). 5μl of running dye (30% glycerol (v/v), 0.1% (w/v) bromophenol blue in running buffer) was added to the DNA sample (0.2μg) and made up to a total volume of 25μl with running buffer. The sample was loaded onto the gel which was run for 4hr at 100mA. The gel was stained for 15 min in a solution of ethidium bromide (0.5μg/ml), viewed on a transilluminator (model C63, U.V. Products) and photographed on Type 57 landfilm on a Polaroid MP3 Land Camera fitted with a 15 and red filters (Wratten).

The mixture was extracted with an equal volume of phenol/chloroform (equilibrated in 50mM Tris/HCl, pH 8.0). The aqueous phase was removed and sequentially extracted with an equal volume of chloroform/isoamyl (24:1) and chloroform. The DNA was precipitated by the addition of 0.1 volumes of 3M sodium acetate, pH 5.5 and 3 volumes of absolute ethanol. The solution was mixed well and placed on a dry-ice/ethanol bath for 30 min. After centrifugation at 13,000g for 5 min the supernatant was discarded and the precipitated pellet dried under a vacuum. The pellet was redissolved in 25μl of water.

b) *Digestion of pDR540 with HindIII and BamH1*

5μg of pDR540 was digested as described in Example 2a). Following organic extraction and precipitation and the DNA was dissolved in 25μl of water.

c) *Ligation of cut pAT153 with cut pDR540*

1μg (5μl) of cut pAT153 from Example 2a) and 1μg (5μl) of cut pDR540 from example 2b) were added to 40μl of ligation mixture containing 62.5mM Tris-HCl, 12.5mM $MgCl_2$, 25mM DTT, 1.25mM ATP, 250μg bovine serum albumin, pH 7.8. The ligation was started by the addition of 1 unit of T4 DNA ligase (1 unit/μl) and incubated for 15 hrs at 15°C.

d) *Bacterial Transformation*

The DNA mixture derived as described in Example 2C) was used to transform *E. coli* JM103.

An overnight culture of *E. coli* JM103 was diluted a hundred-fold into fresh L-broth [30ml - comprises tryptone (10g/l), yeast extract (5g/l) and NaCl (10g/l)]. This culture was grown with vigorous shaking at 37°C. At an OD_{550} of 0.4–0.5 the culture was cooled in an ice-bath for 5 min. The culture was centrifuged at 6,000g for 5 min and the pellet washed with 20ml of ice cold 0.15M NaCl and again centrifuged at 6,000g for 5 min. The pellet was resuspended in 15ml of ice cold solution of $CaCl_2$ (50mM). After 30 min standing on ice the cells were again pelleted by centrifugation and resuspended in 3ml of the $CaCl_2$ solution. To 0.2ml of this cell suspension was added 50μl of a DNA mixture.

Following a further 30 min incubation on ice the suspension was transferred to a waterbath at 43°C and incubated for 2 min. 3ml of L-broth was added and the cells incubated at 37° for 60 min before being plated, at serial dilution, onto selective media. The media was L-agar (L-broth containing 1.5% (w/v) agar) containing ampicillin at 100μg/ml).

e) *Analysis of Hybrid Plasmids*

A small scale plasmid preparation was made of each of the ampicillin resistant transformants as follows:–

1.0ml of culture was grown overnight at 37°C in L-broth. The cells were pelleted by centrifugation at 13,000g for 2 min. The pellet was resuspended in 25μl of 8% (w/v) sucrose, 5% (v/v) Triton X-100 in 50mM Tris-HCl, 50mM EDTA, pH 8.0 and mixed well. 2μl of lysozyme solution (10mg/ml in 250mM Tris-HCl, pH 8.0) was added and the tube placed in a boiling waterbath. After 40 sec the tubes were centrifuged at 13,000g for 10 min. The supernatant (20μl) was removed and mixed with equal volume of isopropanol and placed in a dry ice/ethanol bath. After 10 min the mixture was centrifuged at 13,000g for 10 min.

The supernatant was discarded and the pellet redissolved in 20μl of 0.3M sodium acetate, pH 5.5 and 60μl of ethanol were added. After 30 min in a dry-ice/ethanol bath, the mixture was centrifuged at 13,000g for 5 min, the supernatant discarded and the pellet dried under vacuum. The dried pellet was dissolved in 20μl of 50mM Tris-HCl, pH 8.0 and analysed by agarose gel electrophoresis as described in Example 2a).

Using this analysis a number of transformants were identified which contained hybrid plasmids i.e. a combination of pAT153 and pDR540. Restriction mapping of one such plasmid, GpAT153-Tac, showed it to possess the majority of the plasmid pAT153 (it had lost the 346bp *Hind*III-*Bam*H1 fragment) and the 92 bp fragment containing the *Tac* promoter from plasmid pDR540.

f) *Preparation of GpAT153-Tac*

E. coli JM103 containing GpAT153-Tac were grown in 200ml of L-broth with shaking at 37°C. When an OD_{550} of 0.6 was attained chloramphenicol was added to a final concentration of 170μg/ml. Following a further 16 hour incubation at 37°C the culture was centrifuged at 1000g for 10 min. The bacterial pellet was resuspended in 6ml of 25% (w/v) sucrose in 50mM Tris-HCl, 1mM EDTA, pH 8.0. After cooling on ice, 1.2ml of lysozyme solution (5mg/ml in 250mM Tris-HCl, pH 8.0) was added with gentle mixing. After 5 min incubation at 0°C, 2ml of 250mM

EDTA, pH 8.0 was added, again with gentle mixing. After a further 5 min on ice, 7ml of 0.2% (v/v) triton-X100 in 50mM Tris-HCl, 62.5mM EDTA, pH 8.0 was added with gentle mixing. After a further 15 min on ice, the lysate was centrifuged at 50,000g for 20 mins and the resulting supernatant collected and retained. To 8ml of this supernatant was added 7.6g of solid CsCl and 0.3ml of ethidium bromide solution (10mg/ml).

The mixture was dispensed into two 5ml quick seal tubes and centrifuged at 250,000g for 16 hrs in a Beckmann VTi65 rotor. The plasmid band was observed by illumination with long-wave UV light and removed from the tubes using needle and syringe. The solution was extracted with isoamyl alcohol until colourless. The solution was then heated to 65°C for 15 min (to destroy residual nucleases) and dialysed against 100 volumes of 10mM Tris-HCl, 10mM-NaCl, 0.1mM EDTA pH 8.0 at 4°C. After 4hr this dialysis was repeated.

Plasmid concentration was determined by measuring the OD_{260} of the solution. An OD_{260} value of 1.0 in a path length of 1cm is equivalent to 50μg/ml of DNA.

g) *Digestion of GpAT153-Tac with Sal1 and Ava1*

5μg of GpAT153-Tac was digested as described in Example 2a) except that 10 units of the enzyme *Sal*1 (10 units/μl) and 10 units of the enzyme *Ava*1 (10 units/μl) were substituted for *Bam*H1 and *Hind*III. Also 100mM-NaCl was used rather than 50mM NaCl. Following organic extraction and precipitation the DNA was dissolved in 25μl of water.

h) *Phosphorylation of Oligonucleotides for TrpA Transcription Terminator*

Oligonucleotides 28A and 27B were phosphorylated as described in Example 1f). The oligonucleotides 27A and 28B were not phosphorylated.

i) *Annealing and Ligation of Transcription Terminator*

Oligonucleotides 27A, 28A, 27B and 28B were annealed and ligated as described in Example 1g).

Autoradiography showed that the reaction had proceeded to near completion and a major band corresponding to approximately 80% of the incorporated radioactivity was observed with an estimated length of about 30 base pairs. This ligated fragment was isolated as described in Example 1g) and following ethanol precipitation was dissolved in 20μl of water.

j) *Ligation of cut GpAT153-Tac and Transcription Terminator*

1μg (5μl) of cut GpAT153-Tac from Example 2g) and 5μl of transcription terminator from Example 2i) were ligated as described in Example 2c).

k) *Bacterial Transformation of Vector with Transcription Terminator*

The DNA mixture from Example 2j) was used to transform *E. coli* JM103 as described in Example 2d).

l) *Analysis of Hybrid Plasmids*

Ampicillin resistant transformants were screened as described in Example 2e).

Using this analysis a number of transformants were identified which contained hybrid plasmids i.e. a combination of GpAT153-Tac and the transcription terminator. Restriction mapping one such plasmid, GPAT153-Tac-TT, showed that it possessed the majority of GpAT153-Tac plasmid (it has lost the 774bp Sal1-Ava1 fragment) and the 32bp transcription terminator.

m) *Preparation of GpAT153-Tac-TT*

Plasmid DNA of GpAT153-Tac-TT was prepared as described in Example 2f).

n) *Digestion of GpAT153-Tac-TT with BamH1*

5μg of GpAT153-Tac-TT prepared as described in Example 2m) was digested as described in Example 2a) except that *Hind*III was omitted. Following organic extraction and ethanol precipitation the DNA was dissolved in 25μl of water.

o) *Digestion of pBR325 with Sau3A*

5μg of plasmid pBR325 was digested as described in Example 2a) except that 10 units of enzyme *Sau*3A (10 units/μl) were used instead of *Bam*H1 and *Hind*III. Following organic extraction and ethanol precipitation the DNA was dissolved in 25μl of water.

p) *Ligation of cut GpAT153-Tac-TT with cut pBR325*

1μg (5μl) of cut GpAT153-Tac-TT prepared as described in Example 2n) and 1μg (5μl) of cut pBR325 prepared as described in Example 2o) were ligated as described in Example 2c).

q) *Bacterial Transformation of GpAT153-Tac-TT/pBR325 Ligation Mixture and Analysis of Hybrid Plasmids*

The ligated material from Example 2p) was used to transform *E. coli* JM103 as described in Example 2d) except that the L-agar also contained chloramphenicol at 25μg/ml.

Colonies resistant to both ampicillin and chloramphenicol were screened as described in Example 2e). Using this analysis a number of transformants were identified which contained hybrid plasmids i.e. a combination of GpAT153-Tac-TT and the chloramphenical acetyl transferase (CAT) gene from pBR325.

Restriction mapping of one such plasmid, GpAT153-Tac-CAT-TT, showed that it possessed all of GpAT153-Tac-TT plasmid plus approximately 100 bp of DNA from pBR325 (including the CAT gene).

Example 3

Construction of CAT-ANF Production Vectors

a) *Preparation of GpAT153-Tac-CAT-TT*

Plasmid DNA of GpAT153-Tac-CAT-TT was prepared as described in Example 2f).

b) *Digestion of GpAT153-Tac-CAT-TT with EcoR1 and Sal1*

5μg of GpAT15-Tac-CAT-TT prepared as described in Example 3a) was digested with 10 units of the enzyme *Eco*R1 (10 units/μl) for 5 min at 37°C in a 50μl reaction mixture (final concentration 50mM-Tris-HCl, 10mM $MgCl_2$, 100mM NaCl, pH 7.5). It has previously been established by the method outlined in Example 2a) that these conditions allowed partial digestion of the plasmid such that all the plasmid molecules were cut at least once. Following organic extraction and ethanol precipitation of the DNA a described in Example 2a) it was redissolved in 50μl of the reaction mixture and digested for 3 hr at 37°C with 10 units of the enzyme *Sal*1 (10 units/μl). Following organic extraction and ethanol precipitation the DNA was dissolved in 25μl of water.

c) *Ligation of cut GpAT153-Tac-CAT-TT with ANF-Eco Oligomer*

1μg (5μl) of cut GpAT153-Tac-CAT-TT prepared as described in Example 3b) and 5μl of the oligomer ANF-Eco prepared as described in Example 1h) were ligated as described in Example 2c).

d) *Bacterial Transformation and Analysis of Hybrid Plasmids*

The ligated material from Example 3c) was used to transform *E. coli* JM103 as described in Example 2d).

Colonies resistant to ampicillin were screened as described in Example 2e).

Using this analysis a number of transformants were identified which contained hybrid plasmids. Restriction mapping of one such plasmid, pTCAE11, showed that it possessed the majority of GpAT153-Tac-CAT-TT and the 122bp ANF-Eco oligomer.

e) *Preparation of pTCAN11*

5μg of GpAT153-TAC-CAT-TT prepared as described in Example 3a) was digested as described in Example 2a) except that 10 units of the enzyme *Nco*1 (10 units/μl) and 10 units of the enzyme *Sal*1 (10 units/μl) were used instead of *Hind*III and *Bam*H1 and 100mM NaCl instead of 50mM NaCl. Following organic extraction and ethanol precipitation the DNA was dissolved in 25μl of water.

1μg (5μl) of this cut GpAT153-Tac-CAT-TT and 5μl of the oligomer ANF-Nco prepared as described in Example 1h) were ligated as described in Example 2c) and the mixture used to transform *E. coli* JM103 as described in Example 2d).

Colonies resistant to ampicillin were screened as described in Example 2e) and a number of transformants were identified which contained hybrid plasmids. Restriction mapping of one such plasmid, pTCAN11, showed that it possessed the majority of GpAT153-Tac-CAT-TT and the 121bp ANF-Nco oligomer.

f) *Construction of pTCAS11*

5μg of GpAT153-Tac-CAT was prepared as described in Example 3a) was digested as described in Example 3b) except that 10 units of the enzyme *Sca*1 (10 units/μl) were used instead of *Eco*R1. Following organic extraction and ethanol precipitation the DNA was dissolved in 25μl of water.

1μg (5μl) of this cut GpAT153-Tac-CAT-TT and 5μl of the oligomer ANF-Sca prepared as described in Example 1h) were ligated as described in Example 2c) and the mixture used to transform *E. coli* JM103 as described in Example 2d).

Colonies resistant to ampicillin were screened as described in Example 2e) and a number of transformants were identified which contained hybrid plasmids. Restriction mapping of one such plasmid, pTCAS11, showed that it possessed the majority of GpAT153-Tac-CAT-TT and the 122bp ANF-Sca oligomer.

g) *Preparation of pTCX2*

20μg of GpAT153-Tac-CAT-TT prepared as described in Example 3a) was digested with 1 unit of *Sca1* for 50 min at 37°C in a 100μl reaction mixture (final concentration 6mM Tris-HCl, 100mM-NaCl, 6mM $MgCl_2$, 1mM DTT, 100μg/ml BSA, pH7.4). It had previously been established by the method described in Example 2a) that these conditions allowed partial digestion of the plasmid such that all plasmid molecules were cut at least once. Following organic extraction and ethanol precipitation of the DNA as described in Example 2a), the DNA was redissolved in 50μl of water.

4μg (10μl) of cut plasmid and 10pmoles each of oligonucleotides XX-A and XX-B were ligated in a 50μl reaction mixture (final concentration 50mM-Tris-HCl, 10mM-$MgCl_2$, 20mM DTT, 1mM ATP, 4mg/ml BSA pH7.8). The ligation was started by the addition of 1 unit of T4 DNA ligase (1 unit/μl) and incubated for 15h at 15°C.

The ligated material was used to transform *E. coli* JM103 as described in Example 2d).

Colonies resistant to ampicillin were screened as described in Example 2e) and a number of transformants were identified which contained hybrid plasmids. Restriction mapping of one such

plasmid, PTCX2, showed that it possessed the majority of GpAT153-Tac-CAT-TT and the 15bp oligomer containing novel *Xba1* and *Xho1* restriction sites.

h) *Preparation of pTCAX21*

Plasmid DNA of pTCX2 was prepared as described in Example 2f). 5μg of pTCX2 was digested with 10 units of enzyme *Xbal* (10 units/μl) for 2h at 37°C in 50μl reaction mixture (final concentration 6mM-Tris-HCl, 50mM-NaCl, 6mM-$MgCl_2$, 100μg/ml BSA, pH7.9). Following organic extraction and ethanol precipitation of the DNA as described in Example 2a), the DNA was redissolved in 50μl of reaction mixture and digested for 3h at 37°C with 10 units of the enzyme *Sal1* (10 unit/μl). Following organic extraction and ethanol precipitation the DNA was dissolved in 25μl of water.

1μg (5μl) of cut pTCX2 and 5μl of the oligomer ANF-Xba prepared as described in Example 1g) were ligated as described in Example 2c).

The ligated material was used to transform *E. coli* JM103 as described in Example 2d). Colonies resistant to ampicillin were screened as described in Example 2e). Using this analysis a number of transformants were identified which contained hybrid plasmids. Restriction mapping of one such plasmid, pTCAX21, showed that it possessed the majority of pTCX2 and the 105bp ANF-Xba oligomer.

Example 4

Sequencing of ANF Constructs

a) *Preparation of pTCAE11, pTCAN11 and pTCAS11*

Plasmid DNA was prepared from all three plasmids as described in Example 2f).

b) *Digestion of Plasmids with EcoR1 and Sal1*

5μg each of plasmids pTCAE11, pTCAN11 or pTCAS11 prepared as described in Example 4a) and 5μg of M13 mp8 RF DNA were digested as described in Example 3e) except that 10 units of the enzyme *EcoR1* (10 units/μl) were used instead of *Nco1*. After organic extraction and ethanol precipitation each plasmid was dissolved in 25μl of water.

c) *Ligation of cut M13mp8 DNA with cut CAT-ANF vector*

1μg (5μl) of cut M13mp8 RFDNA prepared as described in Example 4b) and 1μg (5μl) of cut plasmids pTCAE11, pTCAN11 or pTCAS11 prepared as described in Example 4b) were ligated as described in Example 2c).

d) *Transformation of bacteria by M13 hybrid molecules*

The DNA mixtures derived from Example 4c) were used to transform *E. coli* JM103 essentially as described in Example 2d). However after incubating the mixture for 3 min at 43°C, sequential dilution of cells were added to 3ml of soft L-agar at 45°C (0.6% (w/v) agar in L-broth, molten and maintained at 45°C) containing 10μl of 100mM isopropyl β-D-thiogalactopyranoside (IPTG), 50μl of a 2% (w/v) solution of 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) in dimethylformamide and 0.2ml of exponentially growing *E. coli* JM103 (at an OD_{550} of 0.5). Each suspension was poured directly onto L-plates (1.5% (w/v) agar in L-broth) allowed to solidify and incubated at 37° for 24hr.

e) *Identification of Hybrid M13 'Phage*

Colourless 'phage plaques from Example 4d) were screened for the presence of the ANF gene by using the solution hybridisation method (Hobden, A. N., Read, M. J., Dykes, C. W. and Harford, A. *Anal. Biochem.* (1985) *144* 7578). By this method 3 hybrid 'phage, M13mp8-ANFE, M13mp8-ANFN and M13mp8-ANFS, were identified which contained all of the ANF genes.

f) *Confirmation of the Nucleotide Sequences of Cloned ANF genes*

The ANF genes cloned in M13mp8-ANFE, M13mp8-ANFN and M13mp8-ANFS were sequenced using the dideoxynucleotide method (Sanger, F., Nicklin, S. and Coulson, A. R. *Proc. Natl. Acad.l Sci* U.S.A. (1977) *74* 5463–5467) using as templates single stranded DNA derived from the M13 'phages and the universal 15 base primer DNA 5'-AGTCACGACGTTGTA-3'. Using this method the sequences of the ANF genes were confirmed in the hybrid M13mp8 constructs.

Example 5 Fermentation for the Production of CAT-ANF

Strains: The strain, *E. coli* JM103 containing the plasmids pTCAE11, pTCAN11 or pTCAS11 as described in the previous examples.

Inoculum development stages

i) Ampoule

ii) Slope

iii) Shake flask/florence flask

i) Ampoules of the freeze dried strain are prepared from a slope and stored at room temperature.

ii) *Slope stage* 0.2ml Glycerol/Holme (GH) (or other suitable medium) is added to an ampoule of bacterial. The contents are aseptically transferred onto a slope of Glycerol/Holme agar (GHA) containing 100μg/ml ampicillin and incubated overnight (16h) at 37°.

5ml sterile GH medium are added to the slope and 0.1ml portions of the resulting bacterial suspension are used to inoculate further slopes of GHA containing 100μg/ml ampicillin. The slopes are incubated overnight (16h) at 37° an then stored at 40° for a period not greater than 1 month.

iii) *Shake flask/florence flask* 5ml of sterile GH medium (or other suitable medium e.g. LPSG) are added to a slope from ii) above and the resulting suspension of the surface growth is inoculated into a florence flask.

5ml of the slope suspension is used in inoculate 400ml of GH medium in a 2L florence flask. The flask is incubated overnight (16h) on a rotary shaker at 250 rev/min with a 2" throw and the temperature maintained at 37°.

Fermentation Stage The florence flask grown culture is used to supply 400ml (1% v/v) of inoculum of the fermentation stage. This is carried out in a 70L vessel containing 40L of sterile GH medium (or other suitable medium e.g LPSG). The culture is agitated at a speed of 500 rev/min using 3×6 bladed, 12cm diameter turbine impellers. A constant temperature of 37° is maintained and sterile air is supplied at a rate of 40 L/min. At an OD_{550} of 3.0 (in GH medium) or 15.0 (in LSPG medium), IPTG is added to a final concentration of 1mM. After a further 2–3 hr the fermentation is harvested.

Samples are taken throughout the fermentation for dry weight measurements. Filtered broth samples are used to determine the fermentation parameters which include pH, ammonia, and phosphate.

Media used

LPSG	*g/l*
Lab Lemco	20
Peptone	40
NaCl	2
Glycerol	30

Glycerol/Holme

Na_2HPO_4 $12H_2O$	37.5
KH_2 PO_4	6.0
Na_2SO_4	0.5
NH_4 $7H_2O$	0.2
Thiamine	0.5mg
*Trace elements	1ml
Glycerol	30.

pH adjusted to 7.0

*Trace element solution

$CaCl_2.2H_2O$	0.5
$FeCl_2.6H_2O$	16.7
$ZnSO_4.7H_2O$	0.18
$CoCl_2.6H_2O$	0.18
Na.EDTA	20.1
$CuSO_4.5H_2O$	0.16
$MnSO_4.4H_2O$	0.15

Example 6 Purification of ANF

a) *Cell Breakage*

The harvested cells from the fermentation described in Example 5) are resuspended in 50mM sodium phosphate buffer pH 7.2 containing 0.5M NaCl to an OD_{550} of 20. The cells are disrupted by sonication and the resulting lysate is centrifuged at 25000g for 15 mins to give a pellet of cellular debris and a soluble fraction. These fractions are separated.

b) *Purification of Fusion Protein from Soluble Fractions*

Total soluble protein (15mg) prepared as described in Example 6a) from *E. coli* JM103 containg pTCAS11 or pTCAX21 was dialysed into 50mM-Tris-HCl, 0.1mM-2-mercaptoethanol, pH7.8 and applied to a 1ml column of chloramphericol caproate agarose pre-equilibrated with the same buffer. The column was then washed with buffer until the eluate had an absorbance at 280nm of less than 0.05. A solution containing 50mM-Tris-HCl, 0.1mM-2-mercaptoethanol, 1M-NaCl, 5mM-chloramphenicol, pH 7.8 was applied to the column and the eluate collected. Those fractions containing protein as detected by the Bradford method [Bradford, M.M. (1976) Anal. Biochem *72*, 248–254] were pooled and analysed by SDS-polyacrylamide gel electrophoresis. The pooled fractions contained greater than 95% pure fusion protein.

Fusion protein from *E. coli* JM103 containing pTCAE11 or pTCAN11 is purified using standard methods of protein purification such as differential ammonium sulphate precipitation, gel filtration

and ion exchange chromotography.

c) *Purification of CAT-ANF from Cellular Debris Fraction*

Following solubilisation of the CAT-ANF fusion proteins in cellular debris by addition of SDS, urea, Triton X-100, guanidine- HCl or some other suitable agent, the fusion proteins may be purified as described in Example 6b).

d) *Enzymic cleavage of Fusion Proteins*

Fusion protein was purified as described in Example 6b) from cultures of *E. coli* JM103 containing pTCAX21. The protein (0.5–2mg/ml) was dialysed into 100mM-ammonium acetate, pH4.0 and incubated at 37°C for 24 hours with *S. Aureus* V8 protease using an enzyme:substrate ratio of 1:50.

Fusion protein (0.5-mg/ml) containing enterokinase recognition sites were dialysed into 20mM Tris-HCl, 20mM-$CaCl_2$, pH8.0 and digested at 37°C for 6h with enterokinase at an enzyme:substrate ratio of 1:40.

e) *Purification of ANF*

Following enzymic cleavage of fusion protein by *S. Aureus* V8 protease as described in Example 6d), the digest products were separated by reverse phase HPLC on an Aquapore RP300 column (4.6mm × 25cm). Peptides were eluted from the column using a linear gradient of 0.1% (v/v) trifluroacetic acid (TFA), 12% (v/v) acetonitrile (MeCN) to 0.1% (v/v) TFA, 54% (v/v) MeCN and detected by monitoring absorbance at 214nm using the Waters Absorbance detector model 441. A peak eluting at −27% (v/v) MeCN was confirmed by radioimmunoassay (Peninsula Laboratories Europe Ltd) and amino acid analysis (Millipore-Waters Corporation PICO-TA™ system) to be full-length 1–28 ANF.

Digestion products of fusion protein with enterokinase were separated in a similar method to that described above.

d) *Enzymic Cleavage of Fusion Protein*

The CAT-ANF fusion proteins may be cleaved to yield mature ANF by digestion of the fusion protein with enterokinase in 20mM Tris-HCl, 20mM-$CaCl_2$, pH 8.0 at 37°C. A ratio of 40 parts fusion protein to one part enterokinase is used.

e) *Purification of ANF*

The ANF is purified following enzymic cleavage of the fusion protein using a standard procedure such as that described for the purification of human ANF (Kangawa, K. and Matsuo, H. *Biochem. Biophys. Res. Commun.* (1984) *118* 131–139).

CLAIMS

1. A hybrid protein comprising a first polypeptide having human ANF activity and a carrier polypeptide separated from said first polypeptide by a linker polypeptide containing a recognition site for a proteolytic enzyme.

2. A hybrid protein as claimed in claim 1 containing a recognition site for a proteolytic enzyme selected from enterokinase, thrombin, plasmin, collagenase, *Staphylococcal aureus* V8 protease, Factor Xa and endopeptidase lys c.

3. A hybrid protein as claimed in claim 1 or claim 2 wherein said carrier polypeptide comprises the first 73, 173 or 212 amino acid residues of chloramphenicol acetyltransferase.

4. A hybrid protein as claimed in claim 1 substantially as hereinbefore described.

5. A double-stranded polydeoxyribonucleotide coding for a hybrid protein as claimed in any one of claims 1 to 4 wherein the coding strand comprises (i) a sequence coding for the polypeptide having human ANF activity; (ii) fused with the 5′ end of sequence (i), a sequence coding for the linker polypeptide containing a recognition site for a proteolytic enzyme; (iii) fused with the 5′ end of the sequence (ii) a sequence coding for the carrier polypeptide which sequence begins with a translation initiation codon; and (iv) at least one translation stop codon fused with the 3′ end of sequence (i).

6. A double-stranded polydeoxyribonucleotide as claimed in claim 5 containing the sequence of Fig. 1(a) or Fig. 1(b) of the accompanying drawings.

7. A double-stranded polydeoxyribonucleotide substantially as hereinbefore described.

8. A process for preparing a double-stranded polydeoxyribonucleotide as claimed in any one of claims 5 to 7 which comprises assembling apppropriate deoxyribonucleotides or oligodeoxyribonucleotide sections.

9. A process as claimed in claim 8 substantially as hereinbefore described.

10. A vector containing a double-stranded polydeoxyribonucleotide as claimed in anyone of claims 5 to 7.

11. An expression vector as claimed in claim 10 wherein said double-stranded polydeoxyribonucleotide is incorporated at an appropriate site for expression of the composite structural gene.

12. An expression vector as claimed in claim 11 selected from pTCAE 11, pTCAN 11, pTCAS 11 and pTCAX 21.

13. A process for preparing a plasmid as claimed in any one of claims 10 to 12 which comprises ligating in one or more stages the whole of or a portion of a plasmid with a double-

stranded polydeoxyribonucleotide as claimed in any one of claims 5 to 7.

14. A process as claimed in claim 13 substantially as hereinbefore described with reference to Examples 1 to 3.

15. A microorganism transformed by a vector as claimed in any one of claims 10 to 12.

16. A process for preparing a polypeptide having human ANF acitivity which comprises growing transformed cells containing an expression vector as claimed in claim 11 or claim 12 under conditions whereby the composite structural gene coding in part for human ANF is expressed, contacting the hybrid protein thus produced with the appropriate protoelytic enzyme to release said polypeptide and isolating said polypeptide thus released.

17. A process as claimed in claim 16 substantially as hereinbefore described with reference to Examples 5 and 6.

Printed for Her Majesty's Stationery Office by Burgess & Son (Abingdon) Ltd, Dd 8817356, 1987.
Published at The Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained.